State Estimation for Robotics Solution

Thomas
My personal solution of exercises in "State Estimation for Robotics", D. Barfoot.
Post on https://github.com/taohu1994/State_Estimation_for_Robotics_solution
May 2021

1 Chapter 1 Introduction

NULL

2 Chapter 2. Primer on Probability Theory

2.1 2.5.1

Let $\mathbf{u} = [u_1, \dots, u_n]^T$, $\mathbf{v} = [v_1, \dots, v_n]^T$, we have

$$\mathbf{u}^T\mathbf{v} = u_1v_1 + u_2v_2 + \cdot + u_nv_n$$

and

$$\mathbf{v}^T\mathbf{u} = \begin{bmatrix} v_1u_1 & \dots & \dots & v_1u_n \\ \vdots & v_2u_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ v_nu_1 & \dots & \dots & v_nu_n \end{bmatrix}.$$

It is clear that $\mathbf{u}^T\mathbf{v} = \mathrm{tr}(\mathbf{v}^T\mathbf{u})$.

2.2 2.5.2

Recapping the equation (2.29),

$$I(\mathbf{x}, \mathbf{y}) = H(\mathbf{x}) + H(\mathbf{y}) - H(\mathbf{x}, \mathbf{y}).$$

For two independent variables $\mathbf{x}, \mathbf{y}$, we have

$$I(\mathbf{x}, \mathbf{y}) = 0.$$

Then, it is clear that

$$H(\mathbf{x}) + H(\mathbf{y}) = H(\mathbf{x}, \mathbf{y}).$$

2.3 2.5.3

For a Gaussian random variable, $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, Its covariance matrix could be presented

$$\begin{aligned}\boldsymbol{\Sigma} &= E\{(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T\} \\ &= E\{\mathbf{x}\mathbf{x}^T\} - 2E\{\mathbf{x}\boldsymbol{\mu}^T\} + \boldsymbol{\mu}\boldsymbol{\mu}^T \\ &= E\{\mathbf{x}\mathbf{x}^T\} - \boldsymbol{\mu}\boldsymbol{\mu}^T.\end{aligned}$$

We rewrite it as

$$E\{\mathbf{x}\mathbf{x}^T\} = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T.$$

2.4 2.5.4

We have

$$
\begin{aligned}
E[\mathbf{x}] &= \int_{-\infty}^{+\infty} \mathbf{x}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\big)d\mathbf{x} \\
&= \int_{-\infty}^{+\infty} (\mathbf{y}+\boldsymbol{\mu})(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} \\
&= \int_{-\infty}^{+\infty} \mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} + \int_{-\infty}^{+\infty} \boldsymbol{\mu}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y},
\end{aligned}
$$

where we use the substitution $\mathbf{y} = \mathbf{x} - \boldsymbol{\mu}$. The first integral part could be rewritten

$$
\begin{aligned}
&\int_{-\infty}^{+\infty} \mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} \\
&= \int_{-\infty}^{0} \mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} + \int_{0}^{+\infty} \mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} \\
&= -\int_{0}^{+\infty} -\mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d(-\mathbf{y}) + \int_{0}^{+\infty} \mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} \\
&= 0.
\end{aligned}
$$

Therefore, considering the substitution $\mathbf{z} = \sqrt{\frac{1}{2}}\mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{x}$, the mean became

$$
\begin{aligned}
E[\mathbf{x}] &= \boldsymbol{\mu}(2\pi)^{-\frac{N}{2}}2^{\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}}|\mathbf{\Sigma}|^{\frac{1}{2}} \int_{-\infty}^{+\infty} \exp\big(-\mathbf{z}^T\mathbf{z}\big)d\mathbf{z} \\
&= \boldsymbol{\mu}\pi^{-\frac{N}{2}} \int_{-\infty}^{+\infty} \exp\big(-\mathbf{z}^T\mathbf{z}\big)d\mathbf{z} \\
&= \boldsymbol{\mu}.
\end{aligned}
$$

Note that $\int_{-\infty}^{+\infty} \exp(-x^2)dx = \sqrt{\pi}$.

2.5 2.5.5

We present the variance as follow:

$$
\begin{aligned}
Var[\mathbf{x}] &= (2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \int_{-\infty}^{+\infty} (\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu}) \exp\big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\big)d\mathbf{x} \\
&= (2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}}|\mathbf{\Sigma}|^{\frac{1}{2}}2^{\frac{N}{2}}2^N\mathbf{\Sigma} \int_{-\infty}^{+\infty} \mathbf{y}^T\mathbf{y}\exp\big(-\mathbf{y}^T\mathbf{y}\big)d\mathbf{y} \\
&= (\pi)^{-\frac{N}{2}}\mathbf{\Sigma}2^N 2\int_{0}^{+\infty} \frac{1}{2}\mathbf{t}^{\frac{1}{2}} \exp\big(-\mathbf{t}\big)d\mathbf{t} \\
&= \pi^{-\frac{N}{2}}\Gamma(\frac{3}{2})^N 2^N\mathbf{\Sigma} \\
&= \pi^{-\frac{N}{2}}(\frac{\sqrt{\pi}}{2})^N 2^N\mathbf{\Sigma} \\
&= \mathbf{\Sigma}.
\end{aligned}
$$

Here we use the substitutions $\mathbf{y} = \sqrt{\frac{1}{2}}\mathbf{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})$ and $\mathbf{t} = \mathbf{y}^T\mathbf{y}$. Note that

$$\int_0^{+\infty} \mathbf{t}^{\frac{1}{2}} \exp\big(-\mathbf{t}\big)d\mathbf{t} = \Gamma(\frac{3}{2}),$$

where $\Gamma(\cdot)$ is the Gamma function.

2.6 2.5.6

Product of K Gaussian PDFs could be written

$$\begin{aligned}
&\eta\prod_{k=1}^{K}\exp\Big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)^T\Big)\\
&= \eta\exp\big(-\frac{1}{2}\sum_{k=1}^{K}(\mathbf{x}-\boldsymbol{\mu}_k)\mathbf{\Sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)^T\big)\\
&= \eta\exp\Big(-\frac{1}{2}\big(\mathbf{x}^T\sum_{i=1}^{K}\mathbf{\Sigma}_k^{-1}\mathbf{x} - \mathbf{x}^T\sum_{k=1}^{K}\mathbf{\Sigma}_k^{-1}\boldsymbol{\mu}_k - (\sum_{k=1}^{K}\boldsymbol{\mu}_k^T\mathbf{\Sigma}_k^{-1})\mathbf{x} + \sum_{k=1}^{K}\boldsymbol{\mu}_k^T\mathbf{\Sigma}_k^{-1}\boldsymbol{\mu}_k)\Big)\\
&= \eta\exp\Big(-\frac{1}{2}\big(\mathbf{x}^T\mathbf{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu} - \boldsymbol{\mu}^T\mathbf{\Sigma}^{-1}\mathbf{x} + \boldsymbol{\mu}^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}\big)\Big)\exp\big(-\frac{1}{2}(\sum_{k=1}^{K}\boldsymbol{\mu}_k^T\mathbf{\Sigma}_k^{-1}\boldsymbol{\mu}_k - \boldsymbol{\mu}^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu})\big)\\
&= \eta'\exp\big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\big),
\end{aligned}$$

where η' is a normalization constant to enforce the axiom of the total probability, $\mathbf{\Sigma}^{-1} = \sum_{k=1}^{K}\mathbf{\Sigma}_k^{-1}$ and $\mathbf{\Sigma}^{-1}\boldsymbol{\mu} = \sum_{k=1}^{K}\mathbf{\Sigma}_k^{-1}\boldsymbol{\mu}_k$.

2.7 2.5.7

Let $\boldsymbol{\mu}_i, \mathbf{\Sigma}_i$ denote the mean and variance of variable $\mathbf{x}_i$. As $\forall i, j \in [1, K], i \neq j, \mathbf{x}_i, \mathbf{x}_j$ are statistically independent. We have $E[(\mathbf{x}_i - \boldsymbol{\mu}_i)(\mathbf{x}_i - \boldsymbol{\mu}_i)^T] = \mathbf{\Sigma}_i$ and $E[(\mathbf{x}_i - \boldsymbol{\mu}_i)(\mathbf{x}_j - \boldsymbol{\mu}_j)^T] = \mathbf{0}$. The mean of $\mathbf{x}$ could be presented

$$E[\mathbf{x}] = \sum_{(i=1)}^{K}\omega_i E[\mathbf{x}_i] = \sum_{(i=1)}^{K}\omega_i\boldsymbol{\mu}_i = \boldsymbol{\mu}$$

The variance of $\mathbf{x}$ could be presented

$$\begin{aligned}
Var(\mathbf{x}) &= E[\big(\omega_1(\mathbf{x}_1-\boldsymbol{\mu}_1)+\cdots+\omega_K(\mathbf{x}_K-\boldsymbol{\mu}_K)\big)\big(\omega_1(\mathbf{x}_1-\boldsymbol{\mu}_1)+\cdots+\omega_K(\mathbf{x}_K-\boldsymbol{\mu}_K)\big)^T]\\
&= \sum_{i=1}^{K}\omega_i^2 E\{(\mathbf{x}_i-\boldsymbol{\mu}_i)(\mathbf{x}_i-\boldsymbol{\mu}_i)^T\} + \sum_{\forall i,j\in[1,K],i\neq j}\omega_i\omega_j E\{(\mathbf{x}_i-\boldsymbol{\mu}_i)(\mathbf{x}_j-\boldsymbol{\mu}_j)^T\}\\
&= \sum_{i=1}^{K}\omega_i^2 E\{(\mathbf{x}_i-\boldsymbol{\mu}_i)(\mathbf{x}_i-\boldsymbol{\mu}_i)^T\}\\
&= \sum_{i=1}^{K}\omega_i^2\mathbf{\Sigma}_i.
\end{aligned}$$

2.8 2.5.8

Note $\mathbf{x} = [x_1, x_2, \ldots, x_K]^T$. For scalar $x_i \in \mathbf{x}$, we have $x_i \sim \mathcal{N}(0,1)$. And $\forall i, j \in [1, K], i \neq j, E\{x_i x_i\} = 1, E\{x_i x_j\} = 0$. We write the mean of $\mathbf{y}$ as

$$\begin{aligned} E\{\mathbf{x}^T\mathbf{x}\} &= E\{x_1x_1 + \cdots + x_Kx_K\} \\ &= E\{x_1x_1\} + \cdots + E\{x_Kx_K\} \\ &= \sum_{i=1}^{K} E\{x_ix_i\} \\ &= K. \end{aligned}$$

The variance of $\mathbf{y}$,

$$\begin{aligned} \mathrm{Var}(\mathbf{y}) &= E\{(\mathbf{x}^T\mathbf{x} - K)(\mathbf{x}^T\mathbf{x} - K)^T\} \\ &= E\{\mathbf{x}^T\mathbf{x}\mathbf{x}^T\mathbf{x}\} - 2E\{\mathbf{x}\mathbf{x}^T\}K + K^2 \\ &= E\{(x_1x_1 + \cdots + x_Kx_K)(x_1x_1 + \cdots + x_Kx_K)\} - K^2 \\ &= E\{\sum_{i=1}^{K} x_ix_ix_ix_i\} + E\{\sum_{\forall i,j\in[1,K], i\neq j} x_ix_ix_jx_j\} - K^2 \\ &= KE\{x_ix_ix_ix_i\} + (K^2 - K)E\{x_ix_ix_jx_j\} - K^2 \\ &= 2K. \end{aligned}$$

Note that here we use the Isserlis' Theorem and equation (2.40), that

$$E\{x_ix_ix_ix_i\} = 3E\{x_ix_i\}E\{x_ix_i\} = 3.$$

and

$$E\{x_ix_ix_jx_j\} = E\{x_ix_i\}E\{x_jx_j\} + 2E\{x_ix_j\}E\{x_ix_j\} = 1.$$

3 Chapter 3. Linear-Gaussian Estimation

3.1 Exercise 3.6.1

Since we do not have the information about the initial state. We ignore the initial part in the cost function (3.11). Assume the timestep is 5. We have

$$
\begin{aligned}
\mathbf{z} &= [v_1, v_2, v_3, v_4, v_5, y_0, y_1, y_2, y_3, y_4, y_5]^T, \\
\mathbf{x} &= [x_0, x_1, x_2, x_3, x_4, x_5]^T, \\
\mathbf{H} &= \begin{bmatrix}
-1 & 1 & & & & \\
& -1 & 1 & & & \\
& & -1 & 1 & & \\
& & & -1 & 1 & \\
& & & & -1 & 1 \\
1 & & & & & \\
& 1 & & & & \\
& & 1 & & & \\
& & & 1 & & \\
& & & & 1 & \\
& & & & & 1
\end{bmatrix},
\end{aligned}
$$

and

$$
\mathbf{W} = \begin{bmatrix}
Q_1 & & & & & \\
& \ddots & & & & \\
& & Q_5 & & & \\
& & & R_0 & & \\
& & & & \ddots & \\
& & & & & R_5
\end{bmatrix}.
$$

It is obvious that the matrix $\mathbf{H}$ has full column rank. Therefore if $Q \neq 0$ nor $R \neq 0$ there exists a unique solution to the problem.

3.2 Exercise 3.6.2

Substituting Q and R, we have $\mathbf{W} = \mathbf{I}$, then the $\mathbf{H}^T\mathbf{W}^{-1}\mathbf{H}$ could be written

$$
\begin{aligned}
\mathbf{H}^T\mathbf{W}^{-1}\mathbf{H} &= \mathbf{H}^T\mathbf{H} \\
&= \begin{bmatrix}
2 & -1 & 0 & 0 & 0 & 0 \\
-1 & 3 & -1 & 0 & 0 & 0 \\
0 & -1 & 3 & -1 & 0 & 0 \\
0 & 0 & -1 & 3 & -1 & 0 \\
0 & 0 & 0 & -1 & 3 & -1 \\
0 & 0 & 0 & 0 & -1 & 2
\end{bmatrix}.
\end{aligned}
$$

Since the matrix is symmetric positive, we use the Cholesky decomposition. We have $\mathbf{H}^T\mathbf{W}^{-1}\mathbf{H} = \mathbf{L}\mathbf{L}^T$, the matrix $\mathbf{L}$ has the pattern

$$
\mathbf{L} = \begin{bmatrix}
L_{11} & & & & & \\
L_{21} & L_{22} & & & & \\
& L_{32} & L_{33} & & & \\
& & L_{43} & L_{44} & & \\
& & & L_{54} & L_{55} & \\
& & & & L_{65} & L_{66}
\end{bmatrix}.
$$

3.3 Exercise 3.6.3

We have

$$\mathbf{W} = \begin{bmatrix} Q_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & Q_5 & & & & & & \\ & & & R & \frac{R}{2} & \frac{R}{4} & & & \\ & & & \frac{R}{2} & R & \frac{R}{2} & \frac{R}{4} & & \\ & & & \frac{R}{4} & \frac{R}{2} & R & \frac{R}{2} & \frac{R}{4} & \\ & & & & \frac{R}{4} & \frac{R}{2} & R & \frac{R}{2} & \frac{R}{4} \\ & & & & & \frac{R}{4} & \frac{R}{2} & R & \frac{R}{2} \\ & & & & & & \frac{R}{4} & \frac{R}{2} & R \end{bmatrix}.$$

As the sub-matrix

$$R\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{4} & & & \\ \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{4} & & \\ \frac{1}{4} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{4} & \\ & \frac{1}{4} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{4} \\ & & \frac{1}{4} & \frac{1}{2} & 1 & \frac{1}{2} \\ & & & \frac{1}{4} & \frac{1}{2} & 1 \end{bmatrix}$$

has full rank, unless $R = 0$ or $Q = 0$. The matrix $\mathbf{W}$ is invertiable. It is clear that the matrix $\mathbf{H}$ has full column rank, therefore $\mathbf{H}^T\mathbf{W}^{-1}\mathbf{H}$ is invertiable. We conclude that there exists an unique solution for the least square problem when $R \neq 0, Q \neq 0$.

3.4 Exercise 3.6.4

Recapping the equation (3.120), the Kalman filter is

$$\begin{aligned} \check{P}_k &= \hat{P}_{k-1} + Q \\ \check{x}_k &= \hat{x}_{k-1} + v_k \\ K &= \check{P}_k(\check{P}_k + R_k)^{-1} \\ \hat{P}_k &= \big(1 - \check{P}_k(\check{P}_k + R_k)^{-1}\big)\check{P}_k \\ \hat{x}_k &= \check{x}_k + \check{P}_k(\check{P}_k + R)^{-1}(y_k - \check{x}_k) \end{aligned}$$

When $K \to \infty$, we substitute $\check{P}_K$ with $\hat{P}_{K-1} + Q$, the posterior covariance is written

$$\begin{aligned} \hat{P}_K &= \big(1 - (\hat{P}_{K-1} + Q)(\hat{P}_{K-1} + Q + R)^{-1}\big)(\hat{P}_{K-1} + Q) \\ \hat{P}_K &= \frac{R(\hat{P}_{K-1} + Q)}{\hat{P}_{K-1} + Q + R} \\ R\hat{P}_{K-1} &= \hat{P}_K\hat{P}_{K-1} + \hat{P}_K Q + \hat{P}_K R - RQ. \end{aligned}.$$

Since the system is steady, we have $\hat{P} = \hat{P}_K = \hat{P}_{K+1}$. Then the posterior became

$$\hat{P}^2 + \hat{P}Q - RQ = 0.$$

We substitute $\hat{P}$ with $\check{P} - Q$, the prior covariance could be written

$$\check{P}^2 - \check{P}Q - RQ = 0.$$

3.5 Exercise 3.6.5

For the backward Kalman filter, we have the state $\{\hat{\mathbf{x}}_k, \hat{\mathbf{P}}_k, \mathbf{v}_k, \mathbf{y}_{k-1}\}$. Our goal is to estimate the previous step state $\{\hat{\mathbf{x}}_{k-1}, \hat{\mathbf{P}}_{k-1}\}$. The system is reformulated

$$\begin{aligned}
\mathbf{x}_{k-1} &= \mathbf{A}_{k-1}^{-1}\mathbf{x}_k - \mathbf{A}_{k-1}^{-1}\mathbf{v}_k - \mathbf{A}_{k-1}^{-1}\mathbf{w}, \\
\mathbf{y}_{k-1} &= \mathbf{C}_{k-1}\mathbf{x}_{k-1} + \mathbf{n}.
\end{aligned}$$

Therefore, for the prediction step, we have

$$\begin{aligned}
\check{\mathbf{P}}_{k-1} &= \mathbf{A}_{k-1}^{-1}\hat{\mathbf{P}}_k\mathbf{A}_{k-1}^{-T} + \mathbf{A}_{k-1}^{-1}\mathbf{Q}\mathbf{A}_{k-1}^{-T} \\
\check{\mathbf{x}}_{k-1} &= \mathbf{A}_{k-1}^{-1}\hat{x}_k - \mathbf{A}_{k-1}^{-1}\mathbf{v}_k \\
\hat{\mathbf{P}}_{k-1} &= (\check{\mathbf{P}}_{k-1} + \mathbf{C}_{k-1}^T\mathbf{R}^{-1}\mathbf{C}_{k-1})^{-1}.
\end{aligned}$$

We define

$$\mathbf{z} = \begin{bmatrix} \hat{x}_k \\ \mathbf{v}_k \\ \mathbf{y}_{k-1} \end{bmatrix}, \mathbf{x} = \begin{bmatrix} \hat{\mathbf{x}}'_k \\ \hat{\mathbf{x}}_{k-1} \end{bmatrix}, \mathbf{H} = \begin{bmatrix} \mathbf{I} & \\ \mathbf{I} & -\mathbf{A}_{k-1} \\ & \mathbf{C}_{k-1} \end{bmatrix}, \mathbf{W} = \begin{bmatrix} \hat{\mathbf{P}}_k & & \\ & \mathbf{Q} & \\ & & \mathbf{R} \end{bmatrix}.$$

Then the MAP solution is given by $\mathbf{H}^T\mathbf{W}^{-1}\mathbf{H}\hat{\mathbf{x}} = \mathbf{H}^T\mathbf{W}^{-1}\mathbf{z}$. This polynomial could be written,

$$\begin{aligned}
&\begin{bmatrix} \mathbf{P}_k^{-1} + \mathbf{Q}_k^{-1} & -\mathbf{Q}^{-1}\mathbf{A}_{k-1} \\ -\mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1} & \mathbf{A}_{k-1}^T\mathbf{Q}^{-1}\mathbf{A}_{k-1} + \mathbf{C}_{k-1}^T\mathbf{R}_k^{-1}\mathbf{C}_{k-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}'_k \\ \hat{\mathbf{x}}_{k-1} \end{bmatrix} \\
&= \begin{bmatrix} \hat{\mathbf{P}}_k^{-1}\hat{x}_k + \mathbf{Q}^{-1}\mathbf{v}_k \\ -\mathbf{A}_{k-1}^T\mathbf{Q}^{-1}\mathbf{v}_k + \mathbf{C}_{k-1}^T\mathbf{R}_K^{-1}\mathbf{y}_{k-1} \end{bmatrix}
\end{aligned}.$$

Then we eliminate the $\hat{\mathbf{x}}'_k$ by left-multiplying

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1}(\mathbf{P}_k^{-1} + \mathbf{Q}_k^{-1})^{-1} & \mathbf{I} \end{bmatrix}.$$

Now we have

$$\begin{aligned}
&\Big(- \mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1}(\mathbf{P}_k^{-1} + \mathbf{Q}_k^{-1})^{-1}\mathbf{Q}_k^{-1}\mathbf{A}_{k-1} + \mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1}\mathbf{A}_{k-1} + \mathbf{C}_{k-1}^T\mathbf{R}_K^{-1}\mathbf{y}_{k-1}\Big)\hat{\mathbf{x}}_{k-1} \\
&= \mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1}(\mathbf{P}_k^{-1} + \mathbf{Q}_k^{-1})^{-1}(\hat{\mathbf{P}}_k^{-1}\hat{\mathbf{x}}_k + \mathbf{Q}_k^{-1}\mathbf{v}_k) - \mathbf{A}_{k-1}^T\mathbf{Q}^{-1}\mathbf{v}_k + \mathbf{C}_{k-1}^T\mathbf{R}_K^{-1}\mathbf{y}_{k-1}.
\end{aligned}$$

Using the matrix inverse lemma, we have

$$\begin{aligned}
&- \mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1}(\mathbf{P}_k^{-1} + \mathbf{Q}_k^{-1})^{-1}\mathbf{Q}_k^{-1}\mathbf{A}_{k-1} + \mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1}\mathbf{A}_{k-1} = \check{\mathbf{P}}_{k-1}, \\
&- \mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1}(\mathbf{P}_k^{-1} + \mathbf{Q}_k^{-1})^{-1}\mathbf{Q}_k^{-1}\mathbf{A}_{k-1} + \mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1}\mathbf{A}_{k-1} + \mathbf{C}_{k-1}^T\mathbf{R}_K^{-1}\mathbf{y}_{k-1} = \hat{\mathbf{P}}_{k-1}^{-1}, \\
&\mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1}(\mathbf{P}_k^{-1} + \mathbf{Q}_k^{-1})^{-1}\hat{\mathbf{P}}_k^{-1}\hat{\mathbf{x}}_k = \check{\mathbf{P}}_{k-1}^{-1}\mathbf{A}_{k-1}^{-1}\hat{\mathbf{x}}_k, \\
&\mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1}(\mathbf{P}_k^{-1} + \mathbf{Q}_k^{-1})^{-1}\mathbf{Q}_k^{-1}\mathbf{v}_k - \mathbf{A}_{k-1}^T\mathbf{Q}^{-1}\mathbf{v}_k = \check{\mathbf{P}}_{k-1}^{-1}\mathbf{A}_{k-1}^{-1}\mathbf{v}_k, \\
&\mathbf{A}_{k-1}^T\mathbf{Q}_k^{-1}(\mathbf{P}_k^{-1} + \mathbf{Q}_k^{-1})^{-1}(\hat{\mathbf{P}}_k^{-1}\hat{\mathbf{x}}_k + \mathbf{Q}_k^{-1}\mathbf{v}_k) - \mathbf{A}_{k-1}^T\mathbf{Q}^{-1}\mathbf{v}_k = \check{\mathbf{P}}_{k-1}^{-1}\check{\mathbf{x}}_{k-1}.
\end{aligned}$$

Finally we have the corrector

$$\hat{\mathbf{P}}_{k-1}^{-1}\hat{\mathbf{x}}_{k-1} = \check{\mathbf{P}}_{k-1}^{-1}\check{\mathbf{x}}_{k-1} + \mathbf{C}_{k-1}^T\mathbf{R}_k^{-1}\mathbf{y}_{k-1}.$$

We omit the Kalman gain calculation...

3.6 Exercise 3.6.6

Assume

$$\begin{bmatrix} \mathbf{I} & & & & & \\ \mathbf{A} & \mathbf{I} & & & & \\ \mathbf{A}^2 & \mathbf{A} & \mathbf{I} & & & \\ \vdots & \vdots & \vdots & \ddots & & \\ \mathbf{A}^{K-1} & \mathbf{A}^{K-2} & \mathbf{A}^{K-3} & \dots & \mathbf{I} & \\ \mathbf{A}^{K} & \mathbf{A}^{K-1} & \mathbf{A}^{K-2} & \dots & \mathbf{A} & \mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{B}_{1,1} & \dots & \mathbf{B}_{1,K+1} \\ \vdots & \ddots & \vdots \\ \mathbf{B}_{K+1,1} & \dots & \mathbf{B}_{K+1,K+1} \end{bmatrix}$$

. Then all elements of the inverse matrix could be calculated recursively.

$$\begin{aligned} \mathbf{I}\mathbf{B}_{1,1} &= \mathbf{I}; \\ \mathbf{A}\mathbf{B}_{1,1} + \mathbf{B}_{2,1} &= \mathbf{0}; \\ \mathbf{I}\mathbf{B}_{2,2} &= \mathbf{I}; \\ \mathbf{A}^2\mathbf{B}_{1,1} + \mathbf{A}\mathbf{B}_{2,1} + \mathbf{I}\mathbf{B}_{3,1} &= \mathbf{0}; \\ \mathbf{A}\mathbf{B}_{2,2} + \mathbf{I}\mathbf{B}_{3,2} &= \mathbf{0}; \\ \mathbf{I}\mathbf{B}_{3,3} &= \mathbf{I}; \\ &\vdots \\ \mathbf{I}\mathbf{B}_{K+1,K+1} &= \mathbf{I}; \end{aligned}$$

Finally we obtain

$$\mathbf{B} = \begin{bmatrix} \mathbf{I} & & & & & \\ -\mathbf{A} & \mathbf{I} & & & & \\ & -\mathbf{A} & \mathbf{I} & & & \\ & & -\mathbf{A} & \ddots & & \\ & & & \ddots & \mathbf{1} & \\ & & & & -\mathbf{A} & \mathbf{1} \end{bmatrix}.$$

It is clear to verify that

$$\begin{bmatrix} \mathbf{I} & & & & & \\ \mathbf{A} & \mathbf{I} & & & & \\ \mathbf{A}^2 & \mathbf{A} & \mathbf{I} & & & \\ \vdots & \vdots & \vdots & \ddots & & \\ \mathbf{A}^{K-1} & \mathbf{A}^{K-2} & \mathbf{A}^{K-3} & \dots & \mathbf{I} & \\ \mathbf{A}^{K} & \mathbf{A}^{K-1} & \mathbf{A}^{K-2} & \dots & \mathbf{A} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & & & & & \\ -\mathbf{A} & \mathbf{I} & & & & \\ & -\mathbf{A} & \mathbf{I} & & & \\ & & -\mathbf{A} & \ddots & & \\ & & & \ddots & \mathbf{1} & \\ & & & & -\mathbf{A} & \mathbf{1} \end{bmatrix} = \mathbf{I}$$

3.7 Exercise 3.6.7

We notice that the matrix $\mathbf{L}$ has the pattern

$$\mathbf{L} = \begin{bmatrix} L_{11} & & & & & \\ L_{21} & L_{22} & & & & \\ & L_{32} & \ddots & & & \\ & & \ddots & L_{K-1,K-1} & & \\ & & & L_{K,K-1} & L_{K,K} & \\ & & & & L_{K+1,K} & L_{K+1,K+1} \end{bmatrix}.$$

Then, we could compute the elements of its inverse matrix recursively, the computational cost increases linearly with $K + 1$. The inverse computation for $\mathbf{L}$ can be done in $O(N(K + 1))$ time. Since the matrix $\hat{\mathbf{P}}$ is symmetric, the matrix multiplying for $\mathbf{L}^{-1}, \mathbf{L}^{-T}$ is $O(N(\frac{(K+1)^3}{2}))$. Finally, we could conclude the computational cost of computing $\hat{\mathbf{P}}$ is $O(N((K + 1)^3))$.